1984

Redistribution, Inheritance, and Inequality

James B. Davies

10416

ISSN: 0318-725X
ISBN: 0-7714-0582-0

RESEARCH REPORT 8415

REDISTRIBUTION, INHERITANCE, AND INEQUALITY

by

James B. Davies

September, 1984

I. Introduction

It has recently been discovered that a linear tax-transfer system may increase equilibrium inequality in intergenerational models of income distribution (Becker and Tomes, 1979). However, it is not clear why this is the case, or whether the possible disequalizing result is a *curiosum*. It appears widely believed that an increase in inequality due to redistribution is possible in such models because bequests have an intentional "compensatory" element--holding parents' characteristics constant, the desired bequest will be larger the smaller is the "endowed" income of the child. Intuition suggests that if a redistributive scheme taxes inheritances, it may increase inequality by interfering with this equalizing force.

This paper shows that a perverse effect of a linear redistributive scheme is possible in an equilibrium intergenerational model of income distribution, even if bequests are not intentionally compensatory. Even when parents simply save a constant fraction of lifetime income for bequest, inheritances represent an "average" of the resources of previous generations. Given positive but imperfect correlation of earnings across generations, adding such an average to the earnings of the current generation to determine its lifetime income has an equalizing effect. Redistributive schemes which tax inheritances can reduce this equalization by sapping the force of intergenerational accumulation. (This may be true even if the "propensity to bequeath" increases in response to higher taxes.) Whether equilibrium inequality rises or falls depends on whether the possible increase in government transfer payments is large enough to offset this disequalizing effect.

Although this paper confirms that a disequalizing effect of linear redistribution is possible, it provides some hope that it is unlikely to occur in the "leading case". It is widely believed that intertemporal substitution in consumption over the life cycle is inelastic. If the intergenerational analogue--here substitution between parents' consumption and kids' income--is also inelastic the effect of increased taxation on intergenerational accumulation is relatively weak. This reduces the unfavorable impact on the averaging of previous generations' luck in determining inheritances, and tends to give a strong positive effect on transfer payments. Computations with an example indicate that redistribution indeed works when intergenerational substitution is inelastic--except with what appear to be implausible values of some parameters.

The model assumes certainty and a perfect capital market. Factor prices are fixed. The earnings of successive generations are drawn from an exogenous distribution which is stationary except for constant proportional growth. Reproduction is asexual. Finally, each generation maximizes a homothetic function of its own consumption and the income of the next generation.

The paper is organized as follows. Section II sets out the model, and characterizes the equilibrium income distribution. It is found that attempted redistribution has two effects on inequality: one which changes the impact of previous generations' "luck" on the current generation's inheritances, and one which affects the size of transfer payments. These effects are examined in Sections III and IV respectively. Then in Section V an example of the model is provided, with illustrative equilibria. Finally, Section VI compares the results with those obtained in a "naive" model where desired bequest is not affected by kid's endowed income.

II. The Model

We examine a society where reproduction is asexual--a single parent has a single child. A dynasty thus consists of a succession of single individuals. Generations do not overlap.

In the absence of government, a member of generation t would have lifetime income, L_t, composed of earnings and inheritances, E_t and I_t respectively. This income could be expended on consumption, C_t, or bequest, B_t. The wealth inherited by generation t would be determined by the behavior of generation t-1 as follows:

$$(1) \qquad I_t = rB_{t-1} = r(L_{t-1} - C_{t-1}) = r(E_{t-1} + I_{t-1} - C_{t-1})$$

where r is one plus the interest rate.

Simple assumptions are made about factor incomes. The before-tax rate of return to capital, r-1, is exogenous and constant for all time. Earnings are also exogenous. Their size distribution is invariant across generations, except for scale. Mean earnings increase at a constant proportional rate, q-1. Thus the society considered may be viewed as a small open economy in a world experiencing balanced growth.

What form of redistribution does the government impose? We examine a linear tax-transfer scheme based on lifetime income, $E_t + I_t$.

Under this scheme each family receives a basic transfer, or "demogrant", G_t (equal for all families), which together with E_t and I_t, is taxed at the constant proportional rate u. Thus lifetime income before-tax, L_t^b, is given by:

$$(2) \qquad L_t^b = E_t + I_t + G_t$$

and lifetime income after-tax, L_t, is:

(3) $$L_t = (1-u)L_t^b = (1-u)(E_t + I_t + G_t).$$

Using (3), from (1) one's inheritance is:

(4) $$I_t = rB_{t-1} = r(L_{t-1} - C_{t-1}) = r[(1-u)(E_{t-1} + I_{t-1} + G_{t-1}) - C_{t-1}].$$

For the tax-transfer scheme to be self-financing within each generation we must have

(5) $$G_t = u\bar{L}_t^b = u(\bar{E}_t + \bar{I}_t - G_t) = (\frac{u}{1-u})\bar{L}_t.$$

where a bar indicates a mean value. Thus:

(6) $$G_t = (\frac{u}{1-u})(\bar{E}_t + \bar{I}_t)$$

implying a result that simplifies the following analysis considerably:

(7) $$\bar{L}_t = \bar{E}_t + \bar{I}_t .$$

It is assumed that all parents maximize the same utility function, and that this is defined on parent's consumption and child's after-tax lifetime income:

(8) $$U_t = U(C_t, L_{t+1})$$

U is strictly quasi-concave and homothetic.[1] What constraint governs the maximization of (8)? Let the lifetime income of the next generation, conditional on a zero inheritance, be denoted

(9) $$\underline{L}_t = (1-u)(E_t + G_t).$$

Noting that $I_{t+1} = r(L_t - C_t)$, and substituting (9) into the expression for

L_{t+1} implied by (3):

(10) $$L_{t+1} = \underline{L}_{t+1} + r(1-u)(L_t - C_t) \; .$$

We may write (10) in the more familiar form:

(11) $$C_t + \frac{L_{t+1}}{r(1-u)} = L_t + \frac{\underline{L}_{t+1}}{r(1-u)}$$

Expression (11) says that the parent must divide the sum of his own lifetime income and the "endowed" income of his child (referred to below as "family wealth") between his own consumption, C_t, and child's income, L_{t+1}. Given that the interest rate is the same for all, and that U is homothetic and identical across families, all parents will appropriate the same fraction, $(1-\theta)$, of family wealth for their own use:

(12) $$C_t = (1-\theta)\left(L_t + \frac{\underline{L}_{t+1}}{r(1-u)}\right).$$

Substituting (12) into (11):

$$L_{t+1} = r\theta(1-u)L_t + \theta\underline{L}_{t+1}$$

or

(13) $$L_t = \delta\left(L_{t-1} + \frac{\underline{L}_t}{r(1-u)}\right)$$

where $\delta = r\theta(1-u)$. Also we have:

(14) $$I_t = r\theta L_{t-1} - \left(\frac{1-\theta}{1-u}\right)\underline{L}_t$$

From (12) and (13) it is possible to sign $\frac{\partial\theta}{\partial[r(1-u)]}$, and therefore also $\frac{\partial\theta}{\partial r}$ and $\frac{\partial\theta}{\partial u}$. Let the elasticity of substitution between C_t and L_{t+1} be σ.

Then we have:[2]

$$(15)\quad \begin{array}{llll} \text{(i)} & \frac{\partial\theta}{\partial r} < 0,\ \frac{\partial\theta}{\partial u} > 0 & ; & \sigma < 1 \\ \text{(ii)} & \frac{\partial\theta}{\partial r} = \frac{\partial\theta}{\partial u} = 0 & ; & \sigma = 1 \\ \text{(iii)} & \frac{\partial\theta}{\partial r} > 0,\ \frac{\partial\theta}{\partial u} < 0 & ; & \sigma > 1 \end{array}$$

If one wishes, a prior in σ may be formed by analogy to the intertemporal elasticity of substitution in consumption in life-cycle models. This suggests that $\sigma < 1$ is the "leading case".[3] Thus an increase in the tax rate may plausibly be expected to <u>increase</u> the fraction of family wealth passed on to kids.

By repeated substitution, it is easy to show from (13) that:

$$(16)\qquad L_t = \theta \sum_{i=o}^{\infty} \delta^i \underline{L}_{t-i} = \theta \sum_{i=o}^{\infty} \delta^i (1-u)(E_{t-i} + G_{t-i})$$

where $\delta < 1$ is required for convergence. This indicates that proportional differences across families in L_t will be determined by differences in the E_{t-i}'s, and by δ, which governs the lag structure, that is, the relative importance of E_{t-i}'s of different vintage.

If we rewrite (16) as:

$$(17)\qquad L_t = \sum_{i=o}^{\infty} [r(1-u)]^i \theta^{i+1} \underline{L}_{t-i},$$

it is clear that there are two forces governing the benefit received by the current generation from the $\underline{L}_{t-i}$'s (and "luck")[4] of previous generations. First, the higher is the after-tax return on capital, r(1-u), the more one benefits from past accretions to dynastic purchasing power (and the greater is the relative importance of earlier accretions, of course). Second, each generation appropriates a fraction $(1-\theta)$ of any wealth it can get its hands on, leaving θ for the next generation. This reduces the impact of previous $\underline{L}_{t-i}$'s, and in fact must do so sufficiently that $\delta = r(1-u)\theta < 1$ (as noted above). We may clearly obtain very different values of δ, and corresponding degrees of "inheritance" of past resources or luck, depending on the values of $\underline{r}$ and θ.

It is also worth noting from (17) that the impact of $\underline{L}_t$ itself on L_t is reduced by the fraction θ, since the current generation is "taxed" by the previous. This means that $\underline{L}_t$ enters the sum in (17) in much the same way as the $\underline{L}_{t-i}$'s $(i>0)$. There is a difference, however: the interest factor, $r(1-u)$, has no influence on the impact of $\underline{L}_t$ on L_t.

If both E_t's and G_t are growing at the rate q, the value of $\bar{L}_t$ can be easily determined from (16). Taking expectations, and noting that $(\bar{E}_{t-i} + G_{t-i}) = q^{-i}(\bar{E}_t + G_t)$:

$$\bar{L}_t = \theta(1-u)(\bar{E}_t + G_t) \sum_{i=o}^{\infty} \left(\frac{\delta}{q}\right)^i$$

or

$$\bar{L}_t = \left[\frac{\theta(1-u)}{1-\frac{\delta}{q}}\right](\bar{E}_t + G_t) \qquad ; \qquad 0 < \frac{\delta}{q} < 1 \tag{18}$$

(so that L_t will also grow at the steady rate q, ensuring by (5) that the assumption of G_t growing at rate q was correct). Substituting (5) into (18), and rearranging:

$$\bar{L}_t = \frac{\bar{E}_t}{1-\frac{r}{q}+\frac{r(1-\theta)}{\delta}} = \frac{\bar{E}_t}{1-\frac{r}{q}+\frac{1-\theta}{\theta(1-u)}} \tag{19}$$

Taking partial derivatives:

$$(20)\quad \begin{cases} \text{(i)} & \dfrac{\partial \bar{L}_t}{\partial r} = \dfrac{\bar{L}_t^2}{q\bar{E}_t} > 0 \\ \text{(ii)} & \dfrac{\partial \bar{L}_t}{\partial q} = \dfrac{-r\bar{L}_t^2}{q^2\bar{E}_t} < 0 \\ \text{(iii)} & \dfrac{\partial \bar{L}_t}{\partial u} = \dfrac{(\theta-1)\bar{L}_t^2}{\theta(1-u)^2\bar{E}_t} < 0 \\ \text{(iv)} & \dfrac{\partial \bar{L}_t}{\partial \theta} = \dfrac{\bar{L}_t^2}{\theta^2(1-u)\bar{E}_t} > 0 \end{cases}$$

The signs of the partial derivatives in (20) are readily interpreted. The higher r or θ the stronger is the force of intergenerational accumulation, so that higher equilibrium $\bar{L}_t = \bar{E}_t + \bar{I}_t$ is not surprising. Higher u reduces $\bar{L}_t$ by increasing the discount factor applied to kid's endowed income in calculating family wealth (see (11)). Hence parents increase their own consumption, C_t, and reduce their bequests when u rises. (Note that θ is held constant in this analysis.) With r, θ, and (1-u) all exerting the same effect on $\bar{L}_t$ their product, δ, also has a positive impact. Finally, a higher q makes for lower $\bar{L}_t$ since it implies that past $\bar{E}_{t-i}$'s were lower in relation to $\bar{E}_t$. With the accumulation factor δ unchanged, lower earlier resources translate naturally into lower current inheritances, and therefore lower $\bar{L}_t$.

Although $\frac{\partial \bar{L}_t}{\partial u} < 0$ when θ is fixed, it may appear that if θ is allowed to vary, a drop in $\bar{L}_t$ when u rises is not guaranteed. This might appear to follow from:

$$(21)\qquad \frac{d\bar{L}_t}{du} = \overset{(+)}{\frac{\partial\bar{L}_t}{\partial\theta}}\ \overset{(+\text{ or }-)}{\frac{\partial\theta}{\partial u}} + \overset{(-)}{\frac{\partial\bar{L}_t}{\partial u}}$$

As long as $\sigma \geq 1$, $\frac{\partial\theta}{\partial u} \leq 0$ and $\frac{d\bar{L}_t}{du} < 0$ is always obtained. However in what we have suggested is the leading case $(\sigma < 1)$ $\frac{\partial\theta}{\partial u} > 0$ and it might seem possible that $\frac{d\bar{L}_t}{du} > 0$. However, under homotheticity of preferences it turns out that $\bar{L}_t$ must always decline with u.[5]

Turn now to the analysis of distribution:

<u>Lemma 1</u>: L_t and C_t both have stationary distributions except for scale.

<u>Proof</u>: By assumption, E_t has a stationary distribution except for scale. Since G_t and $\bar{E}_t$ grow at the same rate this means that $E_{t-i} + G_{t-i}$ in (16) is also stationary except for scale. Given that θ, δ, and u are constant, from (16) it is clear that L_t has the same property. Finally, from (12) since L_t, and $L_{t+1} = (1-u)(E_{t+1} + G_{t+1})$, are stationary except for scale, so also is C_t.

Q.E.D.

In the discussion of inequality in C_t and L_t, our attention will be confined to the class of scale-independent (zero-degree homogeneous) inequality measures $I = I(Z)$ where I is a function of the continuous distribution of Z ("income") across the population. All indexes I are assumed to obey the Pigou-Dalton "principle of transfers".[6]

Using the homogeneity property, from (16):

$$(22)\qquad I(L_t) = I[\sum_{i=o}^{\infty} \delta^i (E_{t-i} + G_{t-i})]$$

so that inequality in L_t is determined by the lag structure implicit in δ, and the distributions of the $(E_{t-i} + G_{t-i})$'s.

Finally, we have the important result:

$$
\left.\begin{array}{lll} \underline{\text{Lemma 2}}: & \text{(i)} & I(C_t) = I(C_{t'}) = I(C) \\ & \text{(ii)} & I(L_t) = I(L_{t'}) = I(L) \\ \text{and} & \text{(iii)} & I(C) = I(L) \end{array}\right\} \quad -\infty < t, t' < +\infty
$$

<u>Proof</u>: Parts (i) and (ii) follow from the stationarity (except for scale) of C_t and L_t. To see that part (iii) is true, note from (12) and (13):

$$C_{t-1} = \frac{(1-\theta)}{\delta} L_t .$$

That is, although C_t and L_t are not proportional, C_{t-1} and L_t are. Thus $I(C_{t-1}) = I(L_t)$. Using (i) and (ii), $I(C) = I(L)$.

Q.E.D.

This result indicates that in this model, as long as we are interested only in differences within a generation, "inequality" has quite an unambiguous connotation.

III. <u>Effects of Redistribution: I</u>

Equation (16) suggests it may not be difficult to analyze the impact of redistribution in the present model. Current lifetime income is a geometrically-declining distributed lag of past earnings and transfer payments. Since earnings are exogenous there are only two possible influences on the efficacy of redistribution: changes in the lag structure, that is in δ; and movements in the G_{t-i}'s. This section concentrates on the former of these influences. The next section studies under what circumstances the G_{t-i}'s will <u>fall</u> in response to an increase in the tax rate, with a disequalizing effect.

An increase in the tax rate u reduces both $r(1-u)$, and (necessarily), δ.[7] An attempt at redistribution thus has an unambiguous effect on the lag structure in (16). By the corollary to the following theorem this must have a disequalizing effect on the distribution of L_t.

Let Y_{t-i} be a continuous random variable identically distributed for all values of i. Assume that the transition process from Y_{t-i} to Y_{t-i+1} is first-order Markov and non-degenerate. Then we have:

<u>Theorem 1</u>: $I(X^\infty) < I(X'^\infty)$ where

(i) $X^\infty = \sum_{i=o}^{\infty} \beta^i Y_{t-i}$

(ii) $X'^\infty = \sum_{i=o}^{\infty} \beta'^i Y_{t-i}$

(iii) $0 < \beta' < \beta < 1$.

<u>Proof</u>: See Appendix.

<u>Corollary</u>: $I(V^\infty) < I(V'^\infty)$ where

(i) $V^\infty = \sum_{i=o}^{\infty} \delta^i (E_{t-i} + G_{t-i})$

(ii) $V'^\infty = \sum_{i=o}^{\infty} \delta'^i (E_{t-i} + G_{t-i})$

and (iii) $0 < \delta' < \delta < 1$.

<u>Proof</u>: Let $W_{t-i} = q^i(E_{t-i} + G_{t-i}) = q^i \underline{L}_t^b$. Then W_{t-i} is identically distributed for all i, as required of Y_{t-i} in the theorem.[8] We have:

$$\sum_{i=o}^{\infty} \delta^i \underline{L}_{t-i}^b = \sum_{i=o}^{\infty} \left(\frac{\delta}{q}\right)^i q^i \underline{L}_{t-i}^b = \sum_{i=o}^{\infty} \beta^i W_{t-i}$$

where $0 < \beta < 1$. Then, by Theorem 1, for $\beta = \frac{\delta}{q}$ and $\beta' = \frac{\delta'}{q}$ where $0 < \delta' < \delta < 1$:

$$I\left[\sum_{i=o}^{\infty} \beta^i W_{t-i}\right] < I\left[\sum_{i=o}^{\infty} \beta'^i W_{t-i}\right]$$

so that

$$I[\sum_{i=o}^{\infty} \delta^{i}\underline{L}^{b}_{t-i}] < I[\sum_{i=o}^{\infty} \delta'^{i}\underline{L}^{b}_{t-i}]; \qquad 0 < \delta' < \delta < 1.$$

Q.E.D.

The intuition behind this result is that when δ falls, $\Sigma\delta^{i}\underline{L}^{b}_{t-i}$ is more heavily weighted toward recent $\underline{L}^{b}_{t-i}$'s. The $\underline{L}^{b}_{t-i}$'s of more distant generations have less relative influence on the lifetime income of the current generation. To see why this is disequalizing, consider what happens as δ goes to zero. When $\delta = 0$ the only $\underline{L}^{b}_{t-i}$ that counts is $\underline{L}^{b}_{t}$ itself. Letting δ rise above zero "averages" the $\underline{L}^{b}_{t-i}$'s of previous generations with those of the present to determine L_t. If earnings are less than perfectly correlated across the generations this must result in L_t being more equally distributed than $\underline{L}^{b}_{t}$. The reason a fall in δ is disequalizing is that it reduces the averaging that makes $I(L_t) < I(\underline{L}^{b}_{t})$.

Theorem 1 and its corollary can also be used to help predict the effects of differences in the parameters r, θ, and q on equilibrium inequality. Higher values of r or θ raise δ and therefore make for lower inequality via their effect on the lag structure. Since higher θ (greater altruism) must increase $\overline{L}_t$ (see (20)), the higher is θ the larger are government transfer payments (see (5)). Thus higher θ unambiguously leads to less inequality.[9] A weaker argument holds for r. With $\sigma \geq 1$ higher r must also mean higher $\overline{L}_t$ (an increase in r raises θ--see (15)--giving $\overline{L}_t$ a further boost) and therefore an unambiguous drop in inequality. However, with $\sigma < 1$ higher r may lead to lower $\overline{L}_t$, so that this result does not necessarily go through.[10]

In the proof of the corollary to theorem 1 we see that lower q increases β, producing reduced inequality in V^{∞} in just the same way as higher n, θ, or δ. Since lower q also gives higher $\overline{L}_t$ in all cases (and therefore higher transfer payments) it is unambiguous that lower q reduces inequality.

What can we say about the influence of the various parameters on the efficacy of redistribution? We may conjecture, although a rigourous result is not presented here, that the greater is earnings mobility between generations the stronger will be the effect of a fall in δ, and the less equalizing will be redistribution. (The favorable effect of averaging of luck across generations is more important to begin with so its reduction ought to lead to a more significant increase in inequality.)

Finally, the value of σ has a clear impact on the lag structure effect. When $\sigma = 1$ an increase in u will not affect θ, so that the only impact on $\delta = r(1-u)\theta$ is the direct effect. However, with $\sigma > 1$, θ will fall, depressing δ further and making things worse. In the leading case, with $\sigma < 1$ the opposite occurs--there is an increase in θ which partially offsets the effect of increased u on δ. (Recall that δ must always fall as u rises.) Thus the lag structure effect may be least severe in the (leading) $\sigma < 1$ case.

IV. Effects of Redistribution: II

Since a higher tax rate is always disequalizing via its impact on the distributed lag of past $\underline{L}^b_{t-i}$'s in (16), a sufficient condition for redistribution to be disequalizing is that a rise in the marginal tax rate should reduce all G_{t-i}'s--that is, that we should be beyond the peak of the "Laffer Curve". More formally:

Lemma 3: If $\frac{dG_{t-i}}{du} < 0$ for all $i \geq 0$, then $I(L'_t) > I(L_t)$ where:

(i) $L_t = \theta(1-u)\sum_{i=o}^{\infty}\delta^i(E_{t-i} + G_{t-i})$

(ii) $L'_t = \theta'(1-u')\sum_{i=o}^{\infty}\delta'^i(E_{t-i} + G'_{t-i})$

(iii) $u' > u$

and L_t, θ, δ and G_{t-i} are conditional on u, while L'_t, θ', δ', and G'_{t-i} are conditional on u'.

<u>Proof</u>: By the corollary to Theorem 1:

$$(23)\qquad I[\sum_{i=o}^{\infty} \delta'^{i}(E_{t-i} + G_{t-i})] > I[\sum_{i=o}^{\infty} \delta^{i}(E_{t-i} + G_{t-i})] = I(L_t).$$

Now:

$$(24)\qquad \sum_{i=o}^{\infty} \delta'^{i}(E_{t-i} + G'_{t-i}) = \sum_{i=o}^{\infty} \delta'^{i}(E_{t-i} + G_{t-i}) - \sum_{i=o}^{\infty} \delta'^{i}(G_{t-i} - G'_{t-i})$$

If $\frac{dG_{t-i}}{du} < 0$, the second term on the RHS is positive, and the RHS differs from the LHS by an equal absolute subtraction. Hence:

$$I(L'_t) = I[\sum_{i=o}^{\infty} \delta'^{i}(E_{t-i} + G'_{t-i})] > I[\sum_{i=o}^{\infty} \delta'^{i}(E_{t-i} + G_{t-i})]$$

and by (23), $I(L_t) < I(L'_t)$.

Q.E.D.

That $\frac{dG_t}{du} < 0$ should be sufficient for a disequalizing impact of attempted redistribution is hardly surprising: it implies that an increase in u results in <u>taking away</u> an equal absolute amount from all families.

When is $\frac{dG_t}{du} < 0$ likely to occur? From (5)

$$G_t = (\frac{u}{1-u})\bar{L}_t$$

so that the change in G_t depends on the relative response of $\frac{u}{1-u}$ (positive) and $\bar{L}_t$ (negative) to an increase in u. It is clear that the higher σ the more likely is $\frac{dG_t}{du} < 0$, since intergenerational saving will be more responsive to a change in the after-tax rate of return. Thus the "leading case", where

$\sigma < 1$, is once again most conducive to an equalizing effect of redistribution.

Another important influence on $\frac{dG_t}{du}$, and therefore on the efficacy of redistribution is less obvious. From (20), take the second partial derivatives:

$$(25)\quad \begin{aligned} &\text{(i)} \quad \frac{\partial^2 \bar{L}_t}{\partial r \partial u} = \frac{2\bar{L}_t}{q\bar{E}_t} \frac{\partial \bar{L}_t}{\partial u} < 0 \\ &\text{(ii)} \quad \frac{\partial^2 \bar{L}_t}{\partial q \partial u} = - \frac{2r\bar{L}_t}{q\bar{E}_t} \frac{\partial \bar{L}_t}{\partial u} > 0 \end{aligned}$$

This indicates that the downward effect of u on $\bar{L}_t$ (θ constant) will be greater the higher is r or the lower is q. Thus high r or low q should tend to make transfer payments, G_t, increase less with u if $\frac{dG_t}{du}$ is positive, and decrease more if it is negative.

Why should higher r (lower q) make $\frac{\partial \bar{L}_t}{\partial u}$ more strongly negative? The intuition is as follows. From (17) r or q do not affect directly the impact of $\underline{L}_t$ on L_t. However, the higher is r the greater the influence of past endowments on current lifetime income via intergenerational accumulation, and the lower is q the greater were those past endowments--producing the same effect. Now if u is increased, with low r (or high q) there is relatively little direct effect on the L_t's since current $\underline{L}_t$ is of dominant importance and is unaffected. Hence the decline in $\bar{L}_t$ will be relatively small. In contrast, if r is high (q low) the wealth being passed forward from previous generations may bulk very large, so that the component of L_t which is directly reduced when u rises is dominant. Hence an increase in u produces a sharp drop in L_t's and $\bar{L}_t$. The importance of this effect is very clear in the computations of the next section.

V. An Example

This section explores an example of the model described above where the utility function (8) is given the explicit form:

$$(8') \qquad U_t = \begin{cases} \dfrac{C_t^{1-\gamma}}{1-\gamma} + \beta \dfrac{L_{t+1}^{1-\gamma}}{1-\gamma} & , \quad \gamma \neq 1 \\ \ln C_t + \beta \ln L_{t+1} & , \quad \gamma = 1 \end{cases}$$

and the elasticity of substitution between C_t and L_{t+1} is $\sigma = \frac{1}{\gamma}$ and is therefore constant. We also assume that earnings are normally distributed[11] and regress toward the mean across the generations.

It is convenient to express the regression to the mean in earnings in terms of changes in human capital, H_t, related to E_t by:

$$(26) \qquad E_{t-i} = w_{t-i}H_{t-i} = w_t q^{-i} H_{t-i}$$

where w_{t-i}, the wage rate, grows at the constant rate q. If H_{t-i} has a stationary distribution, then that of E_{t-i} will be stationary except for scale. The regression mechanism is:

$$(27) \qquad H_{t-i} = (1-v)\overline{H} + vH_{t-i-1} + \epsilon_{t-i}$$

where ϵ_{t-i} is normally distributed, independent of H_{t-i-1}, and has zero mean and constant variance, $V(\epsilon_{t-i}) = V(\epsilon)$. For stationary variance in the H_{t-i}'s, V(H), we require:

$$(28) \qquad V(\epsilon) = (1-v^2)V(H).$$

Clearly $I(E_{t-i}) = I(E) = I(H_{t-i}) = I(H)$.

Appendix B derives an expression for the equilibrium coefficient of variation, CV(L) = CV(C) in this example.[12] The response of CV(L) to a full menu of tax rates is shown in Table 1 for a "central case" which uses the apparently most plausible parameter values, and for alternative cases which use higher and lower values of four key parameters--r, v^2, β, and σ.

The central case parameters are set as follows. Basic inequality in earnings, CV(E), is set at .340.[13] The growth rate, q, is given the moderate value of 1.01 per annum. (Since there is much less uncertainty about the plausible range of q than about the other parameters, Table 1 does not explore sensitivity to alternative choices of q.) Popular estimates of the mean household rate of return on net worth place r at about 1.04 on an annual basis. (See, for example, Boskin, 1978, p. 19; and Feldstein, Green and Sheshinski, 1978, p. 64.) A generation is assumed to span 25 years, so that the annual q and r of 1.01 and 1.04 respectively correspond to values per generation of 1.28 and 2.67. By analogy to the evidence on the intra-generational intertemporal elasticity of substitution in consumption to which we have already referred, σ is set at 0.5. A best-guess for v^2 appears to be about 0.4.[14] Finally, there is no empirical evidence on β. It was set here by experimenting with alternative values to see which would give a realistic $\frac{\bar{I}_t}{\bar{L}_t}$. Blinder (1974) and Davies (1982) both find that this ratio is quite small--in fact less than 0.1. If the tax rate that corresponds to the real world is 0.2, then $\frac{\bar{I}_t}{\bar{L}_t}$ of about the right size is obtained with $\beta = 0.8$.[15]

Table 1 shows that in the central case inequality falls with u throughout. The first column of Table 2, which gives $\bar{L}_t$, indicates this is somewhat surprising. Increases in u lead to a rapid fall in $\bar{L}_t$--from 1.34 at u = 0, to 1.04 at the moderate u = 0.2, and .67 at u = 0.5. While these declines in $\bar{L}_t$ may appear large, Table 3 indicates that they are not large enough to produce a drop in G_t. Transfers increase steadily as the tax rate rises.[16] This equalizing influence swamps the disequalizing effect of higher u on the lag structure throughout the central case.

In turning to sensitivity testing, note that whatever changes are made in v^2 or β, with $\sigma = 0.5$ the rise in G_t with u remains so strong that inequality always declines in response to an increase in the tax rate. However, raising r to 1.06 results in a dramatic change.[17] In Section IV it was shown that higher r is likely to lead to a less favorable response of G_t to an increase in the tax rate. In Table 3 it can be seen that over a significant range of u values, with r = 1.06 increases in u actually reduce G_t. Thus it is not surprising that the impact of redistribution on inequality switches from uniformly negative, to predominantly positive when r is raised to 1.06.

The effect of changing σ is shown in the last three columns of the tables. Raising σ to .75 it remains true that redistribution is equalizing throughout. However, the effect is much less pronounced: the percent reduction in the CV going from the zero tax case to a rate of 30%, for example, is just 9% compared to 15% in the central case. By the time we have $\sigma = 1$ the equalizing effect from the increase in transfer payments has become sufficiently weak that the attempt at redistribution is disequalizing throughout. Finally, with $\sigma = 1.5$, over the range where the model has a solution transfers decline monotonically with the tax rate (see Table 3) so that a disequalizing impact must be obtained.

These results appear to show that whether redistribution "works" in this model is highly sensitive to parameter values. Although $r = 1.04$, for example, might seem a good guess, we can hardly reject $r = 1.06$ as obviously "unrealistic". Such an interpretation ignores the fact that the cases where redistribution has a perverse effect in these tables all exhibit unrealistic values of $\bar{I}_t/\bar{L}_t$. At the apparently realistic tax rate of 0.2, the $r = 1.06$ case gives a ratio of 3.1, while the $\sigma = 1.5$ case has inheritances eight times as great as lifetime income on average.

To judge better sensitivity to parameterization, we should consider plausible changes in the parameters, r, β, and σ which give $\bar{I}_t/\bar{L}_t$ in the realistic range for moderate tax rates (say between 0 and 0.1 for a tax rate of about 0.2). Table 4 shows the results of such changes.

In Table 4 we first experiment with increases in r. These can yield realistic $\bar{I}_t/\bar{L}_t$ if we simultaneously reduce either β or σ. The first two columns thus show runs with r set at 1.06 and β and σ reduced as required. In both cases we find that redistribution "works" throughout. The same result is not obtained with a high σ, however. The next two columns show that with $\sigma = 1.5$ reducing β or r to get realistic $\bar{I}_t/\bar{L}_t$ does not prevent redistribution from being mainly disequalizing. (The exception is over low values of u when r is reduced.)

In conclusion, if the intergenerational elasticity of substitution σ is well below unity, an equalizing effect of redistribution appears quite robust in the example. However, with $\sigma > 1$ there is an apparently equally robust disequalizing effect.

VI. Comparison with a "Naive Model"

The above model might be criticized on the grounds that the effect of kid's endowed income on bequest is "too strong". It is therefore interesting that the possibility of a perverse effect of redistribution on inequality survives in a model where this effect is completely removed.

Intergenerational models where parents do not take into account their children's endowed earnings ability or rights to government transfers are frequently examined. (See, for example, Stiglitz, 1969 or Blinder, 1974.) For lack of a better term, refer to these as "naive models". Specifically, the utility function:

(29) $$U_t = U(C_t, (1-u)I_{t+1})$$

is often considered. With homothetic U this gives a constant propensity to consume and bequeath out of L_t, so that:

(30) $$C_t = (1-\hat{\theta})L_t$$

and

(31) $$I_{t+1} = r\hat{\theta}L_t$$

From (31) it is clear that bequests have no intentional compensatory element in the naive model. However, there is an unintended compensatory feature. Inheritances are determined by a distributed lag of the resources of past generations, and therefore introduce an "averaging" of luck across generations similar to that found in the main model of this paper. It turns out that by reducing the impact of this averaging redistribution can again increase inequality.

Equation (2) continues to hold in the naive model. Substituting in (31), by back substitution we obtain:

(32) $$L_t = (1-u)\sum_{i=o}^{\infty} \hat{\delta}^i(E_{t-i} + G_{t-i})$$

where $\hat{\delta} = r\hat{\theta}(1-u)$. Formally, (32) differs from (16) in the previous model only in that it is not premultiplied by $\hat{\theta}$. For the assessment of inequality this makes no difference: only the summation term has significance. Hence, formally, the analysis of inequality in the naive model is much the same as in the previous model. So is the analysis of redistribution.

As in the previous model there are two effects of redistribution: one on the lag structure that must be disequalizing, and one on transfer payments which must be analyzed further.

The equilibrium structure of the naive model is considerably easier to work with than that of the previous model. In place of (19) we obtain:

$$\bar{L}_t = \frac{\bar{E}_t}{1 - \frac{r}{q}\hat{\theta}} \tag{33}$$

from which $\frac{d\bar{L}_t}{du} \gtreqless 0$ as $\frac{\partial\hat{\theta}}{\partial u} \gtreqless 0$. The latter inequalities depend on the elasticity of substitution between consumption and child's inheritance, $\hat{\sigma}$, in just the same way that the sign $\frac{\partial\theta}{\partial u}$ depended on σ:

$$\frac{\partial\hat{\theta}}{\partial u} \quad \begin{matrix} > 0 \ , \ \hat{\sigma} < 1 \\ = 0 \ , \ \hat{\sigma} = 1 \\ < 0 \ , \ \hat{\sigma} > 1 \end{matrix} \tag{34}$$

Hence the impact of u on $\bar{L}_t$ is easy to assess. If we can again argue that an elasticity less than unity is most realistic, in the leading case $\frac{\partial\hat{\theta}}{\partial u} > 0$ and $\bar{L}_t$ <u>rises</u> in response to an increase in the tax rate. Thus transfers must increase. Since $\frac{d\bar{L}_t}{du} = 0$ with $\frac{\partial\hat{\theta}}{\partial u} = 0$ we also have a necessary increase in transfers when $\hat{\sigma} = 1$. It is only in the $\hat{\sigma} > 1$ case that it appears possible to operate beyond the peak of the "Laffer curve" in the naive model.

The analysis of $\frac{d\bar{L}_t}{du}$ appears to indicate that redistribution may work a bit better in the naive model. Here we know for sure that we cannot be beyond the peak of the Laffer curve in the leading case. In the previous model this was not true.

A further major difference between the naive and previous models emerges if we note that to predict what is actually observed, their respective parameters would have to yield the same $\bar{L}_t$.

Comparing (19) and (33) we can see that this requires $\hat{\theta} < \theta$.[18] This makes intuitive sense: in the naive model the inheritance is not reduced by a parental "tax" on $\underline{L}_{t+1}$, as in the compensatory model. Hence, to get the same average level of inheritance the propensity to bequeath out of the resources accumulated up to the parent's generation must be smaller.

The fact that $\hat{\theta} < \theta$ for equal $\bar{L}_t$ has important consequences in the analysis of inequality. A lower $\hat{\theta}$ means $\hat{\delta} < \delta$, and a lag structure that places greater relative weight on more recent $E_{t-i} + G_{t-i}$'s. Thus in the naive model the lag structure is less effective in "averaging out" the inequality in $E_t + G_t$. Hence equilibrium inequality will be greater. (With u and $\bar{L}_t$ the same in the two models, the G_{t-i}'s are also the same, so that there is no difference in the impact of transfers on inequality.)

VII. Conclusion

This paper has investigated the impact of attempts at redistribution on equilibrium inequality in a model where bequests are motivated by altruism and parents take into account children's "endowed incomes" when determining intergenerational saving. It has been demonstrated that an unambiguously disequalizing impact of a linear tax-transfer scheme defined on lifetime income is possible. The impact on inequality has two components, one of which always acts in the disequalizing direction. This "lag structure" effect increases inequality because it reduces the averaging of the endowed incomes of previous generations with those of the present to determine current income.

Sometimes offsetting, but sometimes reinforcing this effect, is the impact of higher taxation on transfer payments. It is possible for an increase in the tax rate to result in a reduction in equilibrium government revenue, and therefore in transfer payments. When this occurs the redistributive scheme must be disequalizing. The "leading case", where the intergenerational elasticity of substitution is less than unity, is clearly less likely to produce this result than the elastic case. However, it is not possible to tell from theory alone how "likely" is a perverse redistributional effect.

The theoretical predictions have been confirmed in an example with constant elasticity of substitution between parent's consumption and child's income and a normal distribution of earnings with "regression to the mean". In addition, it is found that except with high interest rates the equalizing effect of larger transfers dominates the disequalizing effect via the lag structure when intergenerational substitution is relatively inelastic. The example thus justifies some optimism that redistribution may typically "work".

Finally, we have comparedthe model with a "naive" version in which parents have homothetic preferences over consumption and bequest. The formal analysis is similar. The same two main effects operate. However, if the two models are to compete in explaining a given observed equilibrium mean level of income, the naive model must assume a lower rate of intergenerational accumulation, that is less averaging of luck across generations, and therefore greater equilibrium inequality. There is also less tendency for transfers to decline when the tax rate is raised. Hence an equalizing effect of redistribution appears more likely.

The comparison with the naive model suggests that whenever inheritance equalizes by averaging past with present luck, redistributive schemes which reduce the strength of intergenerational accumulation will have a disequalizing element. Whether redistribution "works" therefore depends on the strength of its other favorable effects--in the present model a possible increase in government transfer payments. In order to obtain an unambiguously equalizing effect of redistribution it appears that a model where inheritance is a disequalizing institution would have to be formulated. An attractive model incorporating intergenerational altruism which does this has yet to be presented.

Notes

[1]Note that U_t differs considerably from the

$$V_t = V(C_t, C_{t+1}, \dots)$$

assumed in some intergenerational models of income distribution. (See, e.g., Laitner 1979 a and b.) In the present paper no restriction is placed on negative bequests. (These are not allowed in Laitner's work.) When this is the case the U_t formulation may be defended on grounds of greater "realism" despite its asymmetry. With V_t, extremely large negative bequests would be allowed for individuals with much lower earnings than future generations. Also, with U_t knowledge of the complete future stream of E_t's, of (alternatively) modelling of the parents' problem as one of choice under uncertainty, does not have to be incorporated. Becker and Tomes (1979), Ioannides and Sato (1982), and Ioannides (1983) exploit the U_t approach as well.

[2]From (12) and (13)

$$\frac{L_{t+1}}{C_t} = \frac{\delta}{1-\theta} = \frac{r\theta(1-u)}{1-\theta}$$

$$\log\left(\frac{L_{t+1}}{C_t}\right) = \log\theta - \log(1-\theta) + \log[r(1-u)]$$

Since the price of C_t relative to that of L_{t+1} is $r(1-u)$:

$$\sigma = \frac{\partial\log\left(\frac{L_{t+1}}{C_t}\right)}{\partial\log[r(1-u)]} = \frac{\frac{\partial\theta}{\partial[r(1-u)]}}{\theta} + \frac{\frac{\partial\theta}{\partial[r(1-u)]}}{1-\theta} + 1$$

$$= \frac{1}{\theta(1-\theta)}\left\{\frac{\partial\theta}{\partial[r(1-u)]}\right\} + 1$$

Hence:

$$\frac{\partial\theta}{\partial[r(1-u)]} = \theta(1-\theta)(\sigma-1)$$

which implies the signs given in (15), since $1-\theta > 0$.

[3]For summaries of the evidence that $\sigma < 1$ in the life-cycle context, see Davies (1981, pp. 573-4) or Skinner (1983, p. 2). A "best guess" value appears to be $\sigma \approx \frac{1}{2}$.

[4]Becker and Tomes (1979) refer to the random components in endowed income as "luck". Note that the size of each $\underline{L}_{t-i}$ is only partly due to luck--G_{t-i} is the same for all, and E_{t-i} may have a deterministic component (perhaps depending on E_{t-i-1}).

[5]From (12) and (13) $\frac{L_{t+1}}{C_t} = \frac{\delta}{1-\theta}$. By homotheticity, when u rises this ratio must fall. Hence the last term in the denominator of (19) must rise, producing a decline in $\bar{L}_t$.

[6]This principle requires that transfers from richer to poorer should reduce I, while those in the reverse direction should have the opposite effect. (The transfers are assumed not to reverse the relative income rank of donor and donee.) See Sen (1974, p. 27).

[7]If $\sigma \geq 1$, from (15) $\frac{\partial\theta}{\partial u} \leq 0$ so that $\delta = r(1-u)\theta$ will fall when u rises. As noted earlier, from (12) and (13) $\frac{L_{t+1}}{C_t} = \frac{\delta}{1-\theta}$. When u rises, by homotheticity $\frac{L_{t+1}}{C_t}$ must fall. But when $\sigma < 1$, $\frac{\partial\theta}{\partial u} > 0$ and $\frac{\delta}{1-\theta}$ will <u>rise</u> unless δ falls sufficiently. Hence whether the "propensity to bequeath", θ rises or falls, the parameter governing the strength of intergenerational accumulation, δ must always fall in reaction to an increase in u.

[8]Here we assume the G_{t-i}'s are fixed. The next section relaxes this assumption.

[9]For a rigorous statement of this type of argument see the proof of Lemma 3 in the next section.

[10]A case where a rise in r decreases $\bar{L}_t$ is found in the example of Section V. (See the last row in Table 1.) If θ rises sufficiently with r, in the $\sigma < 1$ case, the change in the last term in the denominator of (19) may more than offset the decline produced by the second term becoming a larger negative quantity.

[11]As is well known, observed distributions of earnings are highly skewed, and far better approximated by the lognormal than the normal distribution. In defense of the assumption, skewness in distributions of lifetime earnings is likely less than in annual distributions.

[12]The coefficient of variation is the ratio of the standard deviation to the mean. It is a member of the class of inequality indexes, I. While its properties differ considerably from those of the more popular Gini coefficient (also a member of I), they are not unattractive. See Sen (1974, pp. 27-8).

[13]Davies, St- Hilaire and Whalley (1984) parameterizes earnings mobility over the lifetime according to the results of studies such as Lillard (1977), and obtains CV(E) = .340 for Canadian families, assuming the 1970 situation in Canada represented a steady state.

[14]Blinder (1976, p. 621) argues that the proportion of the variance of earnings explained by "family background" may be viewed as an estimate of v^2, and surveys four studies with average R^2 of .248, suggesting $v^2 \approx 0.25$. Griliches (1979, p. 559) concludes that about 30% of the variance in log earnings is explained by family background in studies using sibling data. Taubman (1976, p. 867) obtains upper and lower bounds of 0.3 and 0.55 for the combined influence of genetics and family environment using data on identical twins. All these studies use annual earnings, producing a downward bias due to transitory earnings. Griliches suggests that correcting this bias could raise R^2 as high as 0.5.

[15] Note that $\beta = 0.8$ is not far from "perfect altruism" (β=1). On an annual basis the implied intergenerational rate of time preference is 0.9%.

[16] In fact for $u \geq 0.3$ successive differences in G_t actually increase, indicating convexity in the "Laffer curve".

[17] Note also that lower mobility (higher v^2) always increases inequality, in line with the conjecture in Section III; greater altruism (higher θ) always reduces equilibrium inequality as the analysis in Section III showed it must; and higher r generally produces a decline in inequality as expected from the previous analysis.

[18] If $\hat{\theta} = \theta$, (33) would be greater than (19) as long as $\delta = r\theta(1-u) = r\hat{\theta}(1-u) < q$. From (18) we know that $\delta < q$ is required for a model solution. (This condition is also required in the "naive model". From (33) $\frac{r}{q}\hat{\theta} < 1$ is required for a solution. This implies $r\hat{\theta}(1-u) = \hat{\delta} < q$.)

Table 1

CV(C) = CV(L) in Example

u	Central Case*	Change in Central Case								
		r =		v^2 =		β =		σ =		
		1.02	1.06	0.3	0.5	0.4	1.2	.75	1.0	1.5
0.0	.219	.268	---	.203	.236	.258	.184	.186	.132	---
0.1	.206	.252	.054	.193	.221	.248	.169	.177	.136	---
0.2	.196	.236	.082	.184	.207	.239	.159	.172	.144	.058
0.3	.186	.221	.099	.176	.196	.230	.151	.170	.151	.107
0.4	.177	.206	.110	.169	.185	.221	.144	.168	.158	.137
0.5	.168	.190	.119	.161	.174	.211	.138	.166	.164	.159
0.6	.159	.174	.124	.153	.164	.200	.130	.164	.169	.179
0.7	.148	.157	.126	.144	.151	.187	.122	.161	.174	.198
0.8	.134	.135	.124	.131	.136	.171	.111	.157	.179	.219
0.9	.111	.106	.113	.110	.112	.143	.093	.148	.184	.247

*$r = 1.04$, $v^2 = 0.4$, $\beta = 0.8$, $\sigma = 0.5$

Note: Missing values indicate the lack of a model solution.

Table 2

$\bar{L}_t$ in Example

u	Central Case*	Change in Central Case: r = 1.02	r = 1.06	v^2 = 0.3	v^2 = 0.5	β = 0.4	β = 1.2	σ = 0.75	σ = 1.0	σ = 1.5
0.0	1.339	.867	---	1.339	1.339	.665	2.429	2.327	5.839	---
0.1	1.183	.813	10.545	1.183	1.183	.609	2.031	1.805	3.224	---
0.2	1.039	.757	4.116	1.039	1.039	.553	1.702	1.418	2.067	8.179
0.3	.907	.698	2.371	.907	.907	.498	1.423	1.119	1.414	2.583
0.4	.783	.637	1.554	.783	.783	.444	1.183	.880	.995	1.305
0.5	.665	.573	1.076	.665	.665	.389	.972	.684	.704	.744
0.6	.553	.504	.760	.553	.553	.333	.783	.519	.489	.435
0.7	.444	.428	.531	.444	.444	.275	.609	.377	.324	.243
0.8	.333	.342	.353	.333	.333	.213	.444	.251	.193	.119
0.9	.213	.235	.201	.213	.213	.141	.275	.134	.088	.039

*$r = 1.04$, $v^2 = 0.4$, $\beta = 0.8$, $\sigma = 0.5$

Note: (a) $\bar{E}_t = 1$.

(b) Missing values indicate the lack of a model solution

Table 3

G_t in Example

u	Central Case*	Change in Central Case								
		r =		v^2 =		β =		σ =		
		1.02	1.06	0.3	0.5	0.4	1.2	0.75	1.0	1.5
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	---
0.1	.131	.090	1.172	.131	.131	.677	.226	.201	.358	---
0.2	.260	.189	1.029	.260	.260	.138	.425	.354	.517	2.045
0.3	.389	.299	1.016	.389	.389	.214	.610	.480	.606	1.107
0.4	.522	.425	1.036	.522	.522	.296	.789	.587	.664	.870
0.5	.665	.573	1.076	.665	.665	.389	.972	.684	.704	.744
0.6	.830	.756	1.140	.830	.830	.499	1.174	.779	.733	.652
0.7	1.035	.999	1.240	1.035	1.035	.642	1.421	.880	.756	.567
0.8	1.332	1.369	1.412	1.332	1.332	.852	1.775	1.005	.774	.474
0.9	1.917	2.118	1.808	1.917	1.917	1.270	2.476	1.210	.788	.350

*$r = 1.04$, $v^2 = 0.4$, $\beta = 0.8$, $\sigma = 0.5$

Note: (a) $\bar{E}_t = 1$.

(b) Missing values indicate the lack of a model solution.

Table 4

Alternative Parameterizations of Example*

	CV(C) = CV(L)				$\bar{L}_t$			
	r = 1.06		σ = 1.5		r = 1.06		σ = 1.5	
u	β = .5	σ = .25	β = .56	r = 1.026	β = .5	σ = .25	β = .56	r = 1.026
0	.176	.195	.179	.219	1.715	1.240	2.610	1.876
.1	.173	.185	.187	.211	1.348	1.123	1.579	1.415
.2	.172	.177	.196	.207	1.076	1.013	1.037	1.068
.3	.172	.168	.205	.206	.866	.909	.706	.800
.4	.171	.159	.215	.209	.697	.809	.484	.588
.5	.170	.150	.224	.213	.557	.713	.327	.419
.6	.168	.140	.234	.219	.437	.618	.213	.284
.7	.163	.127	.245	.227	.333	.523	.128	.176
.8	.156	.110	.260	.240	.238	.422	.066	.092
.9	.138	.083	.279	.261	.145	.307	.022	.032

*Parameter values: r = 1.06, v^2 = 0.4, β = 0.8, and σ = 0.5, except as indicated.

References

Becker, Gary S., and Nigel Tomes. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," Journal of Political Economy 87, no. 6 (December 1979): 1153-89.

Blinder, Alan, S. Toward an Economic Theory of Income Distribution [Cambridge, MA: MIT Press, 1974].

_______. "Inequality and Mobility in the Distribution of Wealth," Kyklos XXIX (1976): 607-638.

Boskin, M. J. "Taxation, Saving, and the Rate of Interest," Journal of Political Economy LXXXVI (April 1978 Supplement): S3-S27.

Davies, James B. "Uncertain Lifetime, Consumption, and Dissaving in Retirement," Journal of Political Economy 89, no. 3 (June 1981): 561-577.

_______. "The Relative Impact of Inheritance and Other Factors on Economic Inequality," Quarterly Journal of Economics 47 (August 1982): 471-498.

_______, France St-Hilaire, and John Whalley. "Some Calculations of Lifetime Tax Incidence," American Economic Review, September, 1984.

Feldstein, M. S., J. Green, and E. Sheshinski. "Inflation and Taxes in a Growing Economy with Debt and Equity Finance," Journal of Political Economy 86, no. 2, pt. 2 (April 1978): S53-S70.

Griliches, Z. "Sibling Models and Data in Economics: Beginnings of a Survey," Journal of Political Economy LXXXVII (October 1979 Supplement): S37-S64

Ioannides, Yannis M. "Heritability of Ability, Intergenerational Transfers and the Distribution of Wealth," School of Management, Boston University, mimeo (February 1983).

_______, and Ryuzo Sato. "A General Equilibrium Theory of the Distribution of Wealth and Intergenerational Transfers," School of Management, Boston University, mimeo (March 1982).

Laitner, J. "Household Bequest Behavior and the National Distribution of Wealth," Review of Economic Studies XLVI, no. 3 (July 1979): 467-484, a.

_______. "Household Bequests, Perfect Expectations, and the National Distribution of Wealth," Econometrica 47, no. 5 (September 1979): 1175-1193, b.

Lillard, L. A. "Inequality: Earnings vs. Human Wealth," American Economic Review (March 1977).

Loury, G. C. "Intergenerational Transfers and the Distribution of Earnings," Econometrica 49, no. 4 (July 1981): 843-868.

Sen A. K. On Economic Inequality [Oxford: Clarendon Press, 1973].

Skinner, Jonathan. "Variable Lifespan and the Intertemporal Sensitivity of Consumption," University of Virginia, mimeo (1983).

Stiglitz, Joseph E. "Distribution of Income and Wealth Among Individuals," Econometrica 37 (1969): 382-397.

Taubman, P. "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins," American Economic Review LXVI (December 1976): 858-870.

Tomes, N. "The Family, Inheritance, and the Intergenerational Transmission of Inequality," Journal of Political Economy 89, no. 5 (October 1981), 928-958, a.

_______. "Inheritance and Inequality Within the Family: Equal Division Among Unequals, Or Do the Poor Get More?" University of Western Ontario working paper, 1981, b.

Appendix A

This appendix proves Theorem 1 given in the text.

First, recall the following assumptions on the Y_{t-1}'s:

(i) all Y_{t-i} have the same density function, $f(\cdot)$:

$$f(Y_t) = f(Y_{t-1}) = \cdots = f(Y_{t-n}) = \cdots$$

(ii) the transition process from Y_{t-i} to Y_{t-i+1} is first-order Markov

(iii) there is always some income mobility. It is not the case that every family draws the same Y_{t-i+1} as Y_{t-i}.

Next we need some definitions. An inequality index $I(Z)$, defined on the (continuous) distribution of "income", Z, is one which obeys the "Lorenz criterion". The Lorenz criterion states that:

(i) $I(Z') > I(Z)$ whenever the Lorenz curve from Z' lies everywhere further from the main diagonal than that for Z,

(ii) $I(Z') < I(Z)$ whenever the Lorenz curve for Z' lies everywhere closer to the main diagonal than that for Z,

and (iii) $I(Z') = I(Z)$ whenever the Lorenz curves for Z' and Z coincide.

Note that since all $I(\cdot)$ obey the Lorenz criterion they also obey the Pigou-Dalton principle of transfers,[1] and have the property of scale-independence since multiplying all "incomes" by a constant does not alter a Lorenz curve.

We now prepare to prove Theorem 1 in the text; that is that inequality in the discounted sum $X^\infty = \sum_{i=o}^{\infty} \beta^i Y_{t-i} (0 < \beta < 1)$ varies inversely with β, according

---

[1]See Sen (1973, pp. 54-5).

to the Lorenz criterion. The strategy is to show first that any linear combination of identically distributed income components is less equally distributed than the components themselves unless drawings are perfectly correlated. For example, inequality in X^{∞} is less than in any Y_{t-i}. We then show that increasing β is equivalent to taking a linear combination of an infinite series of random variables, each with the same distribution as X^{∞}, but imperfectly correlated. It follows that inequality in X^{∞} falls with a rise in β.

<u>Lemma 4</u>: $I(Y^2) < I(Y_{t-i}) = I(Y_{t-j}) = I(Y)$ where $Y^2 = \alpha_o Y_{t-i} + \alpha_1 Y_{t-j}$, $0 < \alpha_o, \alpha_1 < \infty$, and $i \neq j$.

<u>Proof</u>: Note first that if Y_{t-i} and Y_{t-j} were perfectly correlated, $\alpha_o Y_{t-i} + \alpha_1 Y_{t-j}$ would have the same distribution (except for scale) and Lorenz curve as Y_{t-i} or Y_{t-j}. Refer to this unique distribution as that of Y. For each percentile, p, denote the share of the bottom p% of the population in this distribution, s*(p). This share function is given by

$$\text{(A.1)} \qquad s^*(p) = \frac{\int_0^{F^{-1}(p)} f(Y)Y \cdot dY}{\mu}$$

where $F^{-1}(p)$ is the inverse distribution function of Y and μ is the overall mean income.

Alternatively, (A.1) can be written:

$$\text{(A.2)} \qquad s^*(p) = \frac{p \cdot \mu_p^*}{\mu}$$

where μ_p^* denotes the mean income of the bottom p%

If Y_{t-i} and Y_{t-j} are not perfectly correlated the share function is given by:

$$\text{(A.3} \qquad s(p) = \frac{\int_{o}^{F_{Y^2}^{-1}(p)} f_{Y^2}(Y^2)Y^2 dY^2}{\mu}$$

$$= \frac{\int_{o}^{F_{Y^2}^{-1}(p)} f_{Y^2}(Y^2)(\alpha_o Y_{t-i} + \alpha_1 Y_{t-j})dY^2}{\mu}$$

This can be written

$$\text{(A.4)} \qquad s(p) = p \cdot (\alpha_o \phi_p^{Y^2,Y_{t-i}} + \alpha_1 \phi_p^{Y^2,Y_{t-j}})$$

where $\phi_p^{Y^2,Y_{t-i}}$, for example, denotes the mean value of Y_{t-i} drawn for the "bottom p%" of the Y^2 distribution.

Now for any given p it is either the case that Y_{t-i} and Y_{t-j} are perfectly correlated, or not, for the bottom p% of the X distribution. If perfect correlation holds, $s(p) = s^*(p)$ and the Y^2 Lorenz curve coincides with the Y_{t-i} Lorenz curve below p. However, if perfect correlation does not hold either:

$$\phi_p^{Y^2,Y_{t-i}} > \mu_p^*,$$

$$\phi_p^{Y^2,Y_{t-j}} > \mu_p^*,$$

or $$\phi_p^{Y^2,Y_{t-i}}, \phi_p^{Y^2,Y_{t-j}} > \mu_p^* .$$

(If not all drawings have been from the bottom p% of the Y_{t-i} distribution, for example, some weight is transferred from values of $Y_{t-i} < F^{-1}(p)$ to $Y_{t-i} > F^{-1}(p)$, unambiguously increasing the mean Y_{t-i} drawing for the bottom p%

of the Y^2 distribution.) Hence when perfect correlation does not hold $s(p) > s^*(p)$. Thus the Y^2 Lorenz curve may coincide with the Y_{t-i} Lorenz curve over some range (which must begin at the bottom, of course), but will lie above over some range, so that $I(Y^2 = \alpha_o Y_{t-i} + \alpha_1 Y_{t-j}) < I(Y)$ according to the Lorenz partial ordering.

Q.E.D.

<u>Lemma 5</u>: $I(Y^\infty) < I(Y)$ where

(i) $Y^\infty = \sum_{i=o}^{\infty} \alpha_i Y_{t-i}$

(ii) $\sum_{i=o}^{\infty} \alpha_i < \infty$

<u>Proof</u>: Let $Y^n = \sum_{i=o}^{n} \alpha_i Y_{t-i}$. Assume that $I(Y^n) < I(Y)$. Then we will show $I(Y^{n+1}) < I(Y)$ as well.

In a similar procedure to that adopted in the previous proof, let $s^*(p)$ be the share function when Y^n and Y_{t-n-1} are perfectly correlated:

(A.5) $$s^*(p) = \int_o^{F^{-1}_{Y^{n+1}}(p)} f_{Y^{n+1}}(Y^{n+1}) Y^{n+1} dY^{n+1}$$

This may be written:

(A.6) $$s^*(p) = \frac{p \cdot \mu_p^{Y^{n+1}}}{\mu^{Y^{n+1}}}$$

or

$$(A.7) \qquad s^*(p) = \frac{p \cdot [\mu_p^{Y^n} + \alpha_{n+1}\, \mu_p^{Y_{t-n-1}}]}{\mu^{Y^{n+1}}}$$

$$= p \cdot [\omega^{Y^n} \frac{\mu_p^{Y^n}}{\mu^{Y^n}} + \omega^{Y_{t-n-1}} \frac{\mu_p^{Y_{t-n-1}}}{\mu^{Y_{t-n-1}}}]$$

where $\omega^Z = \frac{\mu^Z}{\mu^{Y^{n+1}}}$, $Z = Y^n, Y_{t-n-1}$. Hence

$$(A.8) \qquad s^*(p) = \omega^{Y^n} s^{Y^n}(p) + \omega^{Y_{t-n-1}} s^{Y_{t-n-1}}(p)$$

Now, since ω^{Y^n} and $\omega^{Y_{t-n-1}}$ sum to unity:

$$s^{Y_{t-n-1}}(p) \leq s^*(p) \leq s^{Y^n}(p) \;;\; 0 \leq p \leq 1$$

and $\qquad s^{Y_{t-n-1}}(p) < s^*(p) < s^{Y^n}(p)$ for some p in (0,1).

Hence the Y^{n+1} Lorenz curve lies <u>between</u> the Y_{t-n-1} and Y^n Lorenz curves when drawings from Y^n and Y_{t-n-1} are perfectly correlated.

According to our assumptions on the Y_{t-i}'s we do not, of course, have perfect correlation. By an argument similar to that of the previous proof, there is at least some (upper) range of p values over which $s(p) > s^*(p)$, while there may be a lower range of p values where $s(p) = s^*(p)$. (Once again, if in forming Y^{n+1}'s for the bottom p% of the Y^{n+1} population, any drawing occurs from outside the bottom p% of the Y^n or Y_{t-n-1} distributions $s(p)$ is forced above $s^*(p)$.) Hence the Y^{n+1} Lorenz curve lies above that obtained in case of perfect correlation of Y^n and Y_{t-n-1}. But the latter lies above the Y Lorenz curve. Hence the Y^{n+1} Lorenz curve lies above the Y Lorenz curve. That is, $I(Y^{n+1}) < I(Y)$.

Summing up, we have shown

(i) $I(Y^2 = \alpha_o Y_t + \alpha_1 Y_{t-1}) < I(Y)$

(ii) If $I(Y^n = \sum_{i=o}^{n} \alpha_i Y_{t-i}) < I(Y)$ then $I(Y^{n+1}) < I(Y)$.

By induction, therefore, $I(Y^{n+j}) < I(Y)$ for all integers j such that $0 < j < \infty$, and $I(Y^\infty) < I(Y)$.

Q.E.D.

<u>Theorem 1</u>: $I(X^\infty) < I(X'^\infty)$ where

(i) $X^\infty = \sum_{i=o}^{\infty} \beta^i Y_{t-i}$

(ii) $X'^\infty = \sum_{i=o}^{\infty} \beta'^i Y_{t-i}$

(iii) $0 < \beta' < \beta < 1$

<u>Proof</u>: Let $\beta = k\beta'$, so that $k > 1$.

$$X^\infty = \sum_{i=o}^{\infty} \{\beta'^i Y_{t-i} + [(k\beta')^i Y_{t-i} - \beta'^i Y_{t-i}]\}$$

$$= X'^\infty + \sum_{i=o}^{\infty} (k^i - 1)\beta'^i Y_{t-i}$$

Let $\lambda_i = k^i - 1$.

Now $\lambda_i - \lambda_{i-1} = k^i - k^{i-1}$

$$= k^i(1 - \frac{1}{k})$$

Since $\lambda_o = 0$,

$$\lambda_i = \sum_{j=1}^{i} k^j(1 - \frac{1}{k})$$

Thus

$$X^{\infty} = X'^{\infty} + \sum_{i=1}^{\infty} \{[\sum_{j=1}^{i} k^{j}(1 - \frac{1}{k})]\beta'^{i} \; Y_{t-i}\}$$

$$= X'^{\infty} + (1 - \frac{1}{k})\{k\beta' Y_{t-1} + (k+k^{2})\beta'^{2} Y_{t-2} + (k+k^{2}+k^{3})\beta'^{3} \; Y_{t-3} + \cdots\}$$

$$= X'^{\infty} + (1 - \frac{1}{k}) \sum_{i=1}^{\infty} (k\beta')^{i}[\sum_{j=o}^{\infty} \beta'^{j} \; Y_{t-i-j}]$$

$$= X'^{\infty} + T$$

Note that T is a linear combination of discounted sums $= \sum_{j=o}^{\infty} \beta'^{i} \; Y_{t-i-j}$ all of which have the same distribution and degree of inequality as X'^{∞}, due to the fact that the transition process is first-order Markov. Thus $I(T) < I(X'^{\infty})$, by Lemma 8. By Lemma 5 it is therefore the case that $I(X^{\infty}) = I(X'^{\infty} + T) < I(X'^{\infty})$.

Q.E.D.

Appendix B

This appendix shows how the example of Section V is solved.

Maximizing (8') subject to (11) we obtain a solution in which:

$$\text{(B.1)} \qquad \theta = \frac{1}{1 + \beta^{1/\gamma}[r(1-u)]^{\frac{1-\gamma}{\gamma}}}$$

from which it is straightforward to confirm that (15) holds (recalling that $\sigma = \frac{1}{\gamma}$).

Substituting (26) in (16):

$$\text{(B.2)} \qquad L_t = \theta \sum_{i=o}^{\infty} \delta^i (1-u)(w_{t-i}H_{t-i} + G_{t-i})$$

Now, repeated back substitution in (27) indicates that:

$$\text{(B.3)} \qquad H_{t-i} = \bar{H} + \sum_{j=o}^{\infty} v^j \epsilon_{t-i-j}$$

Noting that $w_{t-i} = w_t q^{-i}$ and $G_{t-i} = G_t q^{-i}$, and substituting (B.3) into (B.2):

$$\text{(B.4)} \qquad L_t = \theta \sum_{i=o}^{\infty} (\frac{\delta}{q})^i (1-u)(w_t\bar{H} + G_t) + \theta \sum_{i=o}^{\infty} (\frac{\delta}{q})^i \tilde{w}_t \sum_{j=o}^{\infty} v^j \epsilon_{t-i-j}$$

where $\tilde{w}_t = (1-u)w_t$.

Since the first term on the RHS is clearly the expectation of L_t:

$$\text{(B.5)} \qquad L_t = \bar{L}_t + \theta\tilde{w}_t \sum_{i=o}^{\infty} (\frac{\delta}{q})^i \sum_{j=o}^{\infty} v^j \epsilon_{t-i-j}$$

This can be rewritten:

$$\text{(B.6)} \qquad L_t \begin{cases} = \bar{L}_t + \theta\tilde{w}_t \sum_{i=o}^{\infty} (\frac{\nu^{i+1} - v^{i+1}}{\nu - v})\epsilon_{t-i} \; ; \; (\nu \neq v) \\ = \bar{L}_t + \theta\tilde{w}_t \sum_{i=o}^{\infty} \nu^i (i+1)\epsilon_{t-i} \; ; \; (\nu = v) \end{cases}$$

where $\nu = \frac{\delta}{q}$.

Assuming $\nu \neq v$, and taking the variance:

$$V(L_t) = \theta^2 \tilde{w}_t^{\,2} \left[\frac{1 + \nu v}{(1-\nu v)(1-\nu^2)}\right] V(H) \tag{B.7}$$

Dividing this by the square of (19) we obtain:

$$CV^2(L_t) = CV^2(L) = \left[\frac{(1+\nu v)\Psi^2}{(1-\nu v)(1-\nu^2)}\right] CV^2(H) \tag{B.8}$$